

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

# 해커톤 프로젝트

DX 파이썬데이터분석 중급 (3차)

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

## 해커톤 Mini PJT Kaggle 주제

- Porto Seguro는 브라질의 자동차 보험 회사 (bias 된 데이터 96:4)
  - Porto Seguro는 브라질의 자동차 보험 회사
  - 이 대회의 목적은 어떤 차주가 내년에 보험 청구를 할 확률을 예측하는 것
  - -1 값이 결측치
  - roc score 로 평가
  - <https://www.kaggle.com/competitions/porto-seguro-safe-driver-prediction>
- 산탄데르 은행 제품 추천
  - 고객의 상황에 가장 적합한 금융 상품 자동 추천
  - 2015-01-28 ~ 2016-06-28 까지 총 1년 6개월사이 월별 데이터
  - 약 1300만개의 대용량 데이터
  - 각 클래스의 Mean Average Precision
  - <https://www.kaggle.com/competitions/santander-product-recommendation>
- Bike
  - 도시 자전거 공유 시스템에 사용되는 자전거 대여량 예측
  - <https://www.kaggle.com/competitions/bike-sharing-demand>

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

## 해커톤 Mini PJT

- 1조:
  - 신현아, 윤덕노, 최환영, 황인수
- 2조:
  - 김동찬, 한재희, 이경수, 김현재
- 3조:
  - 박병건, 임지수, 박지영, 박문규
- 4조:
  - 강용운, 김경수, 박규현, 조재진
- 5조 (개인 프로젝트 세션)
  - 황인수
  - 신동현
  - 오진현
  - 김형석
  - 이용원

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

## Milestone

### • 첫째 날

- 2교시까지 프로젝트 주제 선정
- 4교시까지 프로젝트 진행계획 작성  
(슬라이드에 1장 할애하여 문제정의, 평가지표 설정)
- 6교시까지 데이터 분석 모델 선정
- 8교시까지 데이터 분석 전처리 모듈 구현
- **17:20분까지 데이터 분석서 제출**
  - 메일 제목 : "데이터분석서\_데이터이름\_조\_v0.5.pptx"
  - 데이터 분석을 위해 진행한 코드와 결과 설명

### • 둘째 날

- 4교시까지 데이터 모델 구현
- 6교시까지 데이터 모델 평가 및 보고서 작성
- 그 이후 발표자료 정리 및 발표준비
- **15시 20분까지 최종 결과 발표 ppt 제출하기**
  - 메일 제목 : "데이터분석서\_데이터이름\_조\_v1.0.pptx"
  - 데이터 분석을 위해 진행한 코드와 결과 설명
- 15시 30분 ~ 17시 10분 최종 결과 발표

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

## 프로젝트 발표자료에 포함할 내용

- 문제 정의 (태스크 정의):
  - 어떤 데이터셋을 사용하여 어떤 값을 분류 또는 예측하는지
- 모델의 성능평가 지표 설정:
  - 성능평가 지표의 종류 및 목표치
- 데이터 수집 및 적재
- 데이터 탐색 및 전처리
- 모델의 훈련
- 모델의 평가
- 고찰

# 발표시간 아젠다

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

- 발표시간 아젠다
  - 개인발표는 질의응답포함 7분
  - 조별발표는 질의응답포함 15분
- 3:30 ~ 4:05 개인 프로젝트 발표
- 4:10 ~ 4:25 4조
- 4:25 ~ 4:40 3조
- 4:40 ~ 4:55 2조
- 4:55 ~ 5:10 1조
- 5:10 ~ 5:20 과정 정리
- 위와 같이 진행 예정입니다. 발표 시간에 대해 조정 필요한 분들은 메시지 주세요.

- 깃허브의 해커톤프로젝트 폴더에 해커톤 안내자료 pdf파일로 올려두었으니 참고하세요.
  - 그리고, 3가지 데이터의 대회 URL을 포함시켜두었습니다.
  - 캐글의 해당 URL에 접근하셔서 데이터도 다운로드하고 살펴보신 후 주제 선정 진행 해주세요.
- 해커톤 진행하는 동안 휴식은 조별로 자유롭게 조정하시고 발표자료 제출 시간만 잘 지켜주시면 됩니다.
  - 발표자료는 주피터노트북이 아니라 PPT 파일입니다.
  - 오늘은 0.5버전, 내일은 최종버전 이렇게 두 번만 이메일로 제출해주세요.
- 내일 조별과제 발표자는 마이크를 꼭 준비 후 점검 부탁드립니다.

# THANK YOU!